Ionospheric Data Assimilation of Ground GPS TEC by Use of the Kalman Filter

G. A. Hajj[1,2], B. D. Wilson[1], C. Wang[2], X. Pi[1,2], I. G. Rosen[2]

[1]Jet Propulsion Laboratory, California Institute of Technology,
[2]University of Southern California, Department of Mathematics

Submitted to Radio Science, Dec. 12, 2002

A fully 3-dimensional Global Assimilative Ionospheric Model (GAIM) is currently being developed by a joint University of Southern California and Jet Propulsion Laboratory (JPL) team. To estimate the electron density on a global grid, GAIM uses a first-principles ionospheric physics model and one of two estimation techniques: the Kalman filter and 4DVAR (4-Dimensional VARiational). Because of the large dimension of the state (i.e., electron density on a global 3-D grid), implementation of a full Kalman filter is not computationally feasible. Therefore, we have implemented a band-limited Kalman in which a full time propagation of the covariance is performed but only a portion of the covariance matrix is retained. The retained elements are determined based on assumed physical correlation lengths in the ionosphere. The effectiveness of ground GPS data for specifying the ionosphere is assessed by assimilating slant Total Electron Content (TEC) data from 98 sites into the GAIM Kalman filter and validating the retrieved electron density field against independent measurements. A series of GAIM retrievals are presented and validated by comparisons to JPL's Global Ionospheric Maps (GIM) of vertical TEC (VTEC) and VTEC measurements from TOPEX. A statistical evaluation of GAIM and GIM against TOPEX VTEC indicates that GAIM accuracy is comparable or superior to GIM.

1. Introduction

The increasing reliance of our civilization on space technologies has made it clear that creating a "space weather" monitoring capability that provides timely and accurate space environment observations, specifications, monitoring and forecasting, is essential for the safe operation of various defense and commercial systems. The degree of success in creating such a "space weather" system depends mostly on (1) the ability to obtain *global* and *continuous* measurements related to the space environment and (2) the ability to incorporate these various measurements into a physical model in a self-consistent manner.

The state of monitoring and forecasting space weather today can be compared to that of conventional weather monitoring and forecasting almost half a century ago, when observations were fragmentary in space and time, and means of interpreting them were rudimentary. The global and continuous observations obtained in the lower atmosphere (e.g. from weather satellites and radiosondes), the ability to obtain these observations in a timely manner, and the advances made in global weather modeling and in data assimilation algorithms are the main factors that have brought numerical weather prediction (NWP) models to their current level of success.

On the space environment front, we are witnessing a new era. Significant efforts are being planned to collect further information on solar activity and disturbances in the solar wind and the magnetosphere, and data on the upper atmosphere and ionosphere/plasmasphere are becoming truly global and continuous. A case in point is the Global Positioning System (GPS) in which a global network of over 100 ground receivers and regional networks of hundreds to over one thousand receivers created the unprecedented possibility of producing global maps of vertical total electron content (TEC) and ionospheric irregularities in near real-time updated sub-hourly [Pi et al., 1997; Mannucci et al., 1998]. Moreover, within the next few years the number of flight receivers tracking GPS in a limb-viewing geometry for ionospheric occultations [Leitinger et al., 1997; Hajj et al., 1998; Schreiner et al., 1999] will increase to nearly a dozen, providing an extremely dense global set of horizontal cuts through the ionosphere and allowing for accurate 4D global mapping of electron density [e.g., Hajj et al., 2000]. This data set, along with other data such as UV airglow radiances [e.g., Dymond et al., 2001] from current and future missions, provide a truly unprecedented global 3D coverage of the upper atmosphere and ionosphere.

Our objective in this paper is to describe the development of a Global Assimilative Ionospheric Model (GAIM) capable of assimilating a variety of data types including: (1) slant TEC (the integral of electron density along the transmitter-receiver line-of-sight) measurements from GPS ground receivers, (2) change in TEC measurements taken from a low-Earth orbiter (LEO) tracking GPS satellites at positive and negative elevations (i.e., during GPS-LEO occultations), (3) in situ measurements of electron density, and (4) UV airglow radiances. Similar to neutral atmospheric weather models — which assimilate, solve for, and predict 4D fields (3 spatial and 1 temporal) of the atmospheric state parameters such as temperature, specific humidity and wind — GAIM assimilates, solves for, and predicts the electron density in the ionosphere and some of the underlying driving forces ("drivers") such as production rates, dynamo electric fields, and thermospheric neutral densities, temperatures and winds. In doing this GAIM applies two different techniques: (1) the Kalman filter or some approximation thereof, and (2) a 4-dimensional variational (4DVAR) technique. The former technique is used to solve for the electron density in space and time while assuming the "drivers" to be known. The 4DVAR technique solves for the "drivers" from which the electron density is obtained by solving the ionospheric model equations.

Although the two approaches are currently disjoint, they can be combined in an operational scenario. The 4DVAR technique is described elsewhere [Rosen et al., 2001; Pi et al., 2002]. Other background description of GAIM can also be found in [Hajj et al., 2000; Wang et al., 2002]. In this paper, our focus is on the use of the Kalman filter for estimating the ionospheric electron density state and its implementation. Even though the current GAIM is capable of assimilating a number of data sources as listed above, we limit the scope of this study to assimilating ground TEC measurements from a network of 98 globally distributed stations. In doing so, we are following the general tradition and "wisdom" of the NWP community, which introduces new measurements into numerical weather models only after very careful examination and much evaluation. The reason is that each data set has its own nuances and characteristics, and it could influence the data assimilation output in both positive or negative ways. Therefore, optimal assimilation of any data type requires careful tuning of its error covariance, proper evaluation of the data representativeness errors, examination of the effect of the data on the analysis and its covariance, and examination of the consistency of the assumptions used in the Kalman filter and its solution.

In section 2 we review the formulation of the Kalman filter. In section 3 we discuss some practical considerations related to the full Kalman filter such as memory requirements and number of operations, and introduce the band-limited Kalman filter. In section 4, we describe the USC/JPL GAIM physics model and its solution grid. In section 5 we present examples of ionospheric specifications from GAIM retrievals for May, 22-24, 2002 and validation results against GIM and TOPEX. A conclusion is given in section 6.

2. The Kalman filter

We introduce the following definitions (commonly used in NWP):

x_k^t	The *true state*: a discrete representation of the true ionospheric state (density) at time *k*.
$x_k^a = \left\langle x_k^t / m_k^o, x_k^f \right\rangle$	The *analysis*: an estimate of x_k^t given measurements at time *k,* and a forecast x_k^f
$x_k^f = \left\langle x_k^t / m_{k-1}^o \right\rangle$	The *forecast*: an estimate of x_k^t given measurements up to time $k-1$

The observations m_k^o are assumed to be related linearly to the true state x_k^t through an *observation operator* H_k via the equation

$$m_k^o = H_k x_k^t + \varepsilon_k^o \tag{1}$$

$$\varepsilon_k^o = \varepsilon_k^m + \varepsilon_k^r \tag{2}$$

where ε_k^0 is the observational error which is composed of the *measurement error*, ε_k^m, and a *representativeness error*, ε_k^r. The latter is due to the discretization in time and space of the solution for the ionospheric state [for a description of TEC representativeness error see Hajj et al., 2000]. For TEC measurements, the relation between the observations and the state is already linear. For other measurements such UV radiances, a linearization procedure is first required to bring the relation between perturbations of the state and the measurements into the form of Eq. (1). Similarly, a linearization procedure might be required to relate the true state at time *k*+1 to the true state at time k which can then be written in the form

$$x_{k+1}^t = \Psi_k x_k^t + \varepsilon_k^q \tag{3}$$

where Ψ_k is a *forward model* which can represented in a matrix form and ε_k^q is a *process noise* which reflects our uncertainty in the forward model. A linearization procedure is not required in our case since our dynamical model is already linear as we shall see later.

If M_k, R_k *and* Q_k are used to denote the measurement, representativeness and process noise covariances, respectively. Then the Kalman filter can be summarized by the following set of Equations:

$$x_k^a = x_k^f + K_k\left(m_k^o - H_k x_k^f\right) \tag{4}$$

$$K_k = P_k^f H_k^T \left(H_k P_k^f H_k^T + R_k + M_k\right)^{-1} \tag{5}$$

$$P_k^a = P_k^f - K_k H_k P_k^f \tag{6}$$

$$x_{k+1}^f = \Psi_k x_k^a \tag{7}$$

$$P_{k+1}^f = \Psi_k P_k^a \Psi_k^T + Q_k \tag{8}$$

K is known as the Kalman gain, P^a and P^f are the analysis and forecast covariances, respectively. The vector $\left(m_k^o - H_k x_k^f\right)$ is known as the *innovation* vector, and it represents the observation vector minus the predicted observations based on the forecast state. The Kalman filter was first introduced by [Kalman, 1960] and [Kalman and Bucy, 1961] for linear systems or ordinary differential equations. An overview of the use of the Kalman filter for meteorology is given by [Ghil and Malanotte-Rizzoli, 1991; Daley, 1991].

In the data assimilation process, during a given time step (indexed by k in Eqs. 1-8) the state is assumed to be constant (time steps are taken to be 12 minutes in our analysis below). According to the Kalman formalism, at time t_0 (the center of the first time interval), given a forecast (a priori) state, x_0^f, a forecast state covariance, P_0^f, and a set of observations, m_0^O (collected in the interval t_0 - Δt, t_0 + Δt; Δt=6 minutes in our case) with covariances R_0 and M_0, an improved estimate of the state (x_0^a, *the analysis*) at time 0 can be obtained by adding the innovation vector operated upon by the Kalman gain to the forecast state (Eq. 4). Moreover, because of the inclusion of the data during this time step, the forecast state covariance is reduced by the second term on the RHS of Eq. (6) to give the analysis state covariance at time 0. Using a dynamical model of the ionosphere, we can then propagate the state from the first time step (0 min.) to the next one (12 min.) via the forward model, Ψ_k. (Eq. 7). Similarly, we can propagate the analysis state covariance to the next time step via Eq. 8. The process noise, Q_k, in Eq. 8 reflects our uncertainty in the forward model. The propagated state and covariance serve as the forecast for the next time step (12 min.) and the process repeats recursively.

The process of assimilating data continuously and propagating the model at each time step in the manner described above is formally known as *continuous data assimilation* or four-dimensional data assimilation (not to be confused with 4-DVAR) [Daley, 1991]. In continuous data assimilation, the philosophy is that even if the initial condition and/or the model are imperfect, the accumulation of data will gradually force the model integration to the true ionospheric state. In continuous data assimilation, the analysis at time t_k depends on all observations taken at $t<t_k$. However, it is also possible, in principle, to include measurements taken at at $t>t_k$ when estimating the state at time t_k. This can be accomplished, for example, by use of 4-DVAR [see, e.g., Ghil et al., 1997].

For the Kalman filter to be an unbiased, maximum likelihood, minimum variance estimator, the measurements and state errors need to follow Gaussian statistics and be unbiased. In that case, it is possible to show that the Kalman filter estimator (x_k^a) also minimizes the cost functional [*Bierman*, 1977]

$$J_k = \sum \left(m_k^o - H_k x_k^t\right)\left(R_k + M_k\right)^{-1}\left(m_k^o - H_k x_k^t\right)^T + \left(x_k^t - x_k^f\right)P_k^{f\,-1}\left(x_k^t - x_k^f\right)^T \tag{9}$$

where the sum is over all the measurements during step k. This equality can be used to check the consistency of our assumptions on the magnitude of the state and measurement covariances.

3. Approximations to the Kalman filter

Since one of the main purposes of ionospheric data assimilation is to produce an ionospheric specification or prediction that is useful for space weather applications, timeliness is a key factor for a practical implementation of the Kalman filter. Because of the large dimension of the state (i.e., the number of volume elements or voxels used to represent the ionospheric state which is of order N = 10^5 to 10^6), the full Kalman filter may not be computationally feasible in a timely manner. This is true because of memory storage limitations and the number of computations required. Saving the state covariance in memory requires saving N^2 double precision numbers or 80 to 8000 GByte. On the other hand, updating $P_k^f \rightarrow P_k^a$ (Eq. 6) when assimilating M TEC measurements requires of order $M \times N^2$ operations, where each operation is defined as the evaluation of A = B + C * D, where A, B, C and D are double precision numbers. One operation involves a multiplication, a sum and extracting three numbers from high-speed storage and transferring the results to high-speed storage. (The same covariance update requires M×N operations when assimilating in situ measurements.) Furthermore, updating $P_k^a \rightarrow P_{k+1}^f$ (Eq. 8) requires of order N^3 operations. However, the latter transformation (Eq. 8) can be made of order $c_1 N^2$ operations, where c_1 is roughly constant, by taking advantage of the fact that diffusion takes place along magnetic flux tubes and using a common grid to solve for the dynamical equations of the ionosphere and to represent the ionospheric state as described later. In our implementation of GAIM and for a time step of 12 minutes, c_1 is of order 1000.

To appreciate the level of computations needed to perform the full Kalman, consider the following example. A subset of 98 GPS stations from the continuously operational global network operated by the International GPS Service (IGS) collects nearly 700 5-minute-averaged line-of-sight TEC measurements every 12 minutes. Assimilating these measurements and updating the state covariance every 12 minutes requires of order $(700+1000) \times N^2$ operations. An Intel chip with 2GHz speed at best performs only 2×10^9 operations (as defined above) per second. For $N=10^5$-10^6 and for a whole day run, this translates into 2×10^{15}-2×10^{17} operations or 12-1200 days. High resolution operational numerical weather prediction models solve for $\sim 10^7$ variables (Temperature, water vapor, zonal and meridional wind components on a 1 deg. × 1 deg. grid at 30 or more pressure levels). This makes it clear why the full Kalman filter is prohibitive, even on the fastest parallel computers available to date. This is also why the meteorological community has devised numerous approximations to the Kalman filter including optimal interpolation [e.g., Lorenc, 1981], reduced Kalman and band-limited Kalman. For the purpose of this study, we will only consider the band-limited Kalman and we limit the state dimensions to order 10^4 voxels.

In the band-limited approximation, all the Kalman steps (Eqs. 4-8) are performed as usual. However, the state covariance is truncated such that for a given voxel *i*, only a subset of the entire set of voxels will have non-zero correlation (i.e., $P_{ij} \neq 0$ for some preselected $j \in [1,N]$). In the simplest implementation of a band-limited Kalman, a given voxel *i* will have non-zero covariance only with voxels within a specified "correlation volume" as depicted in Fig. 1.

It is worth noting that the name 'band-limited' is strictly valid only for a one dimensional system where the elements can be indexed such that the state covariance matrix is zero everywhere outside a band around the diagonal. However, for higher dimensional systems, the covariance matrix will always have non-zero elements away from the diagonal terms (Fig. 2). Therefore,

when referring to a band-limited Kalman, the term is used only in some abstract sense where in fact the actual state covariance is full, albeit very sparse.

The band-limited Kalman reduces the number of operations required to update the covariance matrix from N^2 to A×N, where A is the number of voxels within the correlation volume, V_{corr}. For a correlation volume of a fixed size, A will grow linearly with N; however, the band-limited Kalman will still reduce the number of operations by a factor V_{corr}/V, where V is the total volume of the modeled region. This ratio can be of order 1/1000 or smaller making the band-limited Kalman very manageable. A realistic representation of the state covariance is paramount for obtaining accurate estimates of the state, especially when the data are sparse relative to the size of the state. The band-limited Kalman filter maintains a sensible covariance while at the same time it reduces the number of computational steps substantially, thereby making it usable for global, medium-resolution, ionospheric runs.

4. Forward Model

A detailed description of the GAIM physical model is given in [Pi et al., 2002]. In summary, we solve the conservation of mass and momentum equations for a plasma, which account for production, loss, transport of the major ionization specie in the F-region (O^+), neutral wind, and electric forcing. These equations can be written as

$$\frac{\partial n}{\partial t} + \nabla \cdot (n\mathbf{V}) = (P - L) \tag{10}$$

$$-\nabla(nk_B T) + nM\mathbf{g} + cn(\mathbf{E} + V \times \mathbf{B}) - nM\nu(V - \mathbf{U}) = 0 \tag{11}$$

where n is the ion number density; $\mathbf{V}$ is velocity; P and L are production and loss rates, respectively; k_B is Boltzmann's constant; T is temperature, M is molecular mass, $\mathbf{g}$ is gravitational acceleration, $\mathbf{E}$ and $\mathbf{B}$ are electric and magnetic fields respectively; ν is the collision frequency for momentum transfer between the atomic oxygen ion and the neutral particles; and $\mathbf{U}$ is neutral wind. An equation similar to Eq. (11) can be obtained for the electrons, and after ignoring terms that are multiplied by the electron's mass, we obtain

$$-\nabla(n_e k_B T_e) - en_e(\mathbf{E} + V_e \times \mathbf{B}) = 0 \tag{12}$$

In addition we also have

$$n_e = n, \qquad n_e V_e = nV \tag{13}$$

The ion and electron densities are obtained by solving the above equations and making use of the empirical or parameterized models of the thermosphere (MSIS) [Hedin et al., 1991], thermospheric winds (HWM) [Hedin et al., 1996], solar EUV (SERF2, [Tobiska, 1991]), and electric fields [e.g., Heppner and Maynard, 1987; Fejer and Scherliess, 1997]. Given all the driving forces, it should be clear that Eq. (11-13) are linear in the ion and electron densities. This linearity is broken once more ions are introduced or the conservation of energy equation is added.

Traditionally, these dynamical equations are rewritten in a moving Lagrangian coordinate frame [e.g., Bailey *et al*., 1993]. The motion of this coordinate frame is dictated by the plasma drift perpendicular to the geomagnetic field lines. This approach introduces significant computational efficiency by transforming a time-dependent partial differential equation in a 3-dimensional

space into a family of time-dependent ordinary differential equations in a 1-dimensional space following the moving flux tubes. However, this approach also introduces significant complications for data assimilation, since the measurements are taken in 3-D space across different field lines (e.g., TEC from ground-to-satellite or satellite-to-satellite links), making the mapping between data and the model parameter space (Eq. 1) very difficult to construct. This, in principle, can be overcome by using two frames: A Lagrangian frame used to solve the dynamical Equations (10-13) and an Eulerian frame where a set of voxels, fixed in space and time, are used to solve for the Kalman filter equations (1-8). Ion and electron densities in the two frames can be related to each other by means of interpolation. We refer to this approach as the dual-frame approach.

A more elegant and efficient approach is to solve both the Kalman filter equations and the dynamical equations in the same Eulerian frame. In this case the volume elements used to discretize the dynamical equations and to perform the Kalman filter are the same and they are defined by the intersection of constant magnetic field lines, constant magnetic potential lines, and constant magnetic longitudes. We refer to this approach as the single-frame approach.

The USC/JPL GAIM uses the single-frame approach where the Earth's magnetic field is modeled by an eccentric tilted dipole (Fig. 3). There are two main advantages to the single-frame over the dual-frame approach. (1) The dual-frame approach requires interpolation of the densities back and forth between the two frames at each time update in the Kalman filter. (2) The time update matrix $\left[\Psi_k\right]_{ij}$ (used in Eqs. 3,7,8) is by definition equal to the partial derivatives $\partial n_i(k+1)/\partial n_j(k)$, where $n_i(k+1)$ *and* $n_j(k)$ are the densities in voxel i at time $k+1$ and voxel j at time k, respectively. In the single-frame approach, it is possible to mathematically construct this matrix of partial derivatives and directly compute it. In the dual-frame approach, additional complications arise in trying to relate the partials of the densities in the Lagrangian frame to those in the Eulerian frame. In the face of these complications, one might have to construct $\left[\Psi_k\right]_{ij}$ by perturbing $n_j(k)$, solving the dynamical forward equations to propagate the state to time $k+1$ and then computing the change in $n_i(k+1)$. This has to be done for one voxel at a time, therefore, requiring as many forward runs as the dimension of the state. Since the single-frame approach uses the same voxels to solve the dynamical and Kalman equations, it is possible to explicitly form the matrix Ψ_k without having to run the forward model. This represent a substantial time saving and makes the implementation of the Kalman filter more feasible.

5. Results and Validation

The USC/JPL GAIM model is now able to assimilate four major data types: absolute slant TEC measurements from ground GPS receivers, change in TEC data from GPS occultations, in situ electron density measurements, and UV airglow radiances. Here we present results from GAIM runs when assimilating only ground GPS-based TEC measurements using the band-limited Kalman filter for the period May 22-24, 2002. This period was chosen to assess the performance of GAIM on quiet days (22 and 24) and a disturbed day (23). For each day, nearly 200,000 GPS TEC measurements, sampled at 1 measurement every 5 minutes, were available from 98 GPS receiver sites using an elevation cutoff of 10 degrees (Figure 4). These TEC measurements are based on dual-frequency phase measurements leveled to the pseudo-range with the GPS instrumental biases determined using the JPL Global Ionospheric Map (GIM) technique [Mannucci, et al., 1998]. A cross section at a constant geomagnetic longitude of the grid used in

the GAIM run is shown in Figure 3. Table 1 summarizes the boundary of the region used in the GAIM run, the vertical and horizontal resolution of the grid and the correlation lengths used in each of the radial, longitudinal and latitudinal directions. Table 2 summarizes the geomagnetic conditions for each day. All three days assumed the same E×B drift climatology (that of June-solar maximum conditions), MSIS for the neutral densities and temperatures, and HWM for neutral wind. In presenting our results below we distinguish between two different GAIM runs: (1) GAIM *climatology* which refers to the GAIM 3-D densities obtained by running the GAIM model without assimilating any data; and (2) GAIM *analysis* which refers to the GAIM 3-D densities obtained by assimilating the ground TEC data described above. In both cases, GAIM yielded a 3-D specification of electron density every 12 minutes for the entire 3 days considered. These can be integrated vertically to create global 2-D maps of vertical TEC (VTEC).

The GAIM analyses of electron density were validated in two ways: (1) comparison of the GAIM VTEC maps to the Global Ionospheric Maps (GIMs) computed from the same ground GPS TEC data, and (2) comparison of GAIM VTEC values to independent TOPEX measurements.

Validation against GIM

GIM is a mapping technique which assumes a thin shell ionospheric model at 450 km. The details of the technique are described in [Mannucci et al., 1998]. In a nutshell, GIM solves for VTEC using a basis set of bi-cubic splines with local support on a spherical shell. The 2D spherical grid is fixed in solar-magnetic coordinates (magnetic local time). By mapping line-of-sight TEC measurements to VTEC at the ionospheric shell piercing point, GIM solves for VTEC on the grid using a square root information filter (SRIF) [Bierman et al., 1977] which is equivalent to the Kalman filter. GIM does not make use of any dynamical model and therefore it is entirely data driven. In regions where there is no data (e.g., gaps in Figure 4), GIM relies on persistence in time to obtain a solution for VTEC. (More specifically, the a priori VTEC value at a given vertex is set equal to its value from the previous time step with a covariance that grows according to a first-order Gauss-Markov process.) Since the stations are rotating underneath the solar-magnetic reference frame in which the grid is defined, nearly all vertices will have some links going through them during the span of a few hours.

Since GIM is a straightforward interpolation of the GPS TEC data using a 2-D shell model, it serves as a proxy for the information content of the GPS dataset. GIM matches the TEC data (mapped to vertical) quite well near the GPS sites. However, GIM interpolation is less accurate at distances greater than 1000 km from the nearest site. Moreover, because of the thin shell model used by GIM, horizontal structures in the ionosphere can potentially create artifacts in the VTEC maps therefore reducing their accuracy near strong gradient regions. Both of these limitations need to be remembered as we compare GAIM to GIM.

To perform the comparison, a GIM global map is updated every 15 minutes for the May 22-24, 2002 period and interpolated to the 12-minute GAIM runs for the same period. Due to space limitation, we only show the comparison between GAIM and GIM at one time frame, about the middle of the period considered (May 23, 1100 UTC), which exhibits features that are representative of all the other time frames.

Figure 5.a shows snapshots of VTEC from the GAIM climatology and GIM, along with maps of the absolute and relative differences. The figure illustrates that the GAIM VTEC level before data assimilation differs by more than 50% from reality (or at least the GIM proxy) in certain regions. By contrast, differences between GAIM analysis and GIM as seen in Figure 5.b are significantly reduced, indicating that the TEC data are being used effectively by GAIM. Further examination of the GAIM analysis shows that GAIM reveals the equatorial anomaly more clearly and with higher resolution than GIM (compare left two panels of 5.b). This is an indication that the limitation induced by the GIM's thin shell model is reduced or eliminated by the 3-D GAIM grid. In addition, since GAIM climatology and GAIM analysis appear to be quite different, even in regions where data is sparse (e.g., in equatorial regions over the Atlantic ocean and Africa), we conclude that the dynamics introduced by the physical modeling of GAIM plays a significant role in the data assimilation.

To appreciate the significance of the role of the physical dynamics used in GAIM, consider the following. Currently, the GAIM Kalman filter only solves for the ion and electron densities and does not adjust any of the drivers. (An extended Kalman filter, where both ion densities and the drivers are estimated, is currently under development and will the subject of a future study.) Therefore, it is expected that if data stops flowing into GAIM, the GAIM analysis will revert to the GAIM climatology on time scales ranging from minutes to several hours. The wide range of time scales is due to the fact that the production, loss, convection, and diffusion of ions respond to the various driving forces on different time scales. For example, the fast recombination rate of molecular ions causes the F1 region to disappear quickly at night, while the slower recombination rate of atomic ions causes the F2 region to last long after dusk when the radiation from the Sun stops. Even when the driving forces are only approximately correct, the dynamical model plays an important role in assigning the correct time scale at different regions and local times in the ionosphere. Effectively, when the time scale is long, the initial condition of the ion densities will have a stronger effect on the evolution of the ionosphere, therefore extending the influence of data over larger regions. When the time scale is short, the initial condition of the ion densities will have little effect on the evolution of the ionosphere, therefore limiting the influence of data to smaller regions. Thus, the model plays a crucial role in assigning the proper time correlation length at different local times, heights and latitudes. This information is completely lost if one uses a constant time scale everywhere, as would be the case if no dynamical model is used to map the state or the state covariance (Eqs. 7 and 8).

Time evolution during a magnetic storm

The May period chosen for our analysis was centered around a magnetic storm. Figure 6 shows the hourly Dst for three days starting at 00:00 UT, May 22, 2002 and indicates the onset of a magnetic storm at 12:00 UT, May 23. The three-hour ap and Kp indices for May 23 are given in Table 3. Figure 7 shows hourly GAIM analysis of VTEC for the 23rd at UT = 12:00, 13:00, …, 17:00, a period which corresponds to the main phase of the storm. For comparisons we also show the corresponding VTEC maps at the same local time for the 22nd as a proxy for the expected ionospheric features during a quiet period. Comparing the two days, a clear enhancement of VTEC at the southern geomagnetic equatorial region is seen at 14:00 UT during the storm. Furthermore, at 15:00 UT, an enhancement of the equatorial anomaly extending from 8:00 to 20:00 local times, can be seen during the storm day relative to the quiet day. This is presumably caused by a storm induced enhancement of the eastward electric field. Figure 8

shows cross sections of GAIM analysis of electron densities at the same universal times as in Figure 7 but only for May 23. The cross sections are taken at 7.5 geographic longitude and therefore the local time is given by the UT + 30 minutes. The features seen in the VTEC maps, such as the enhancement of VTEC at 14:00 UT are clearly seen in the densities as well.

Validation against TOPEX

VTEC below the TOPEX track derived from the dual-frequency altimeter has been used extensively as an independent data source for validation [Ho et al., 1997; Codrescu et al., 2001; Orus et al, 2002]. Given the precision, latitudinal coverage and long time-series of the TOPEX data, it offers a unique and powerful means of validation. When compared to VTEC derived from GPS ground data, TOPEX VTEC are especially challenging given that they are measurements taken exclusively over the ocean where few GPS stations exist. Also, validation against TOPEX will mainly tell us how well GAIM can retrieve VTEC but does not validate the retrieval of densities. Validation of densities will require a different data source and will be the subject of a future investigation.

We start by comparing VTEC from TOPEX, GAIM climatology, GAIM analysis and GIM. Figure 8 shows VTEC for 8 out of the 27 TOPEX tracks on May 23, 2002. Also shown in Figure 8 are the tracks of the TOPEX footprint and neighboring GPS stations used in the assimilation. We note the following features in the comparisons:

1- GAIM climatology matches TOPEX very well in some tracks (e.g., track 10) while it differs significantly in others (e.g., 13, 14, 15, 16, 19).
2- The GAIM analysis is significantly different from the GAIM climatology and compares much better with TOPEX (visible in all tracks) indicating that GPS TEC data is being assimilated effectively.
3- The agreement between the GAIM analysis and TOPEX is quite good in many cases (e.g., 8, 14, 15, 16, 19).
4- Whenever an equatorial anomaly appears in TOPEX, it also appears in the GAIM analysis, (e.g., 9, 13, 19), in some cases with great fidelity (e.g., 19).
5- GAIM appears to be able to capture steep VTEC gradients associated with the equatorial anomaly better than GIM (e.g., 9, 15, 16, 19). This is presumably due to the thin shell limitation of GIM.

Statistical comparison to TOPEX

To further assess the performance of GAIM, we examine histograms of VTEC difference between GAIM climatology and TOPEX (left panels of Figure 9), GAIM analysis and TOPEX (middle panels) and GIM and TOPEX (right panels) for all TOPEX tracks during May 22-24, 2002. Statistical summaries of the histograms are given in table 4. We emphasize that the VTEC differences between GAIM analysis and TOPEX has a standard deviation of $\sigma_{GAIM/A}$ = 5.2 TECU (1 TEC Unit = 10^{16} e/m^2) over the three days which is almost three times better than the standard deviation for GAIM climatology ($\sigma_{GAIM/C}$ = 13.8 TECU), twice better than IRI (σ_{IRI} = 9.6 TECU), and slightly superior to GIM (σ_{GIM} = 5.6 TECU). Both GAIM analysis and GIM VTEC are biased low by 1-2 TECU relative to TOPEX when it should be high given that TOPEX is at 1330 km and GPS is at 20,000 km altitude. One TECU translates into altimetric

range delay of ~2 mm which well within the error budget of TOPEX, therefore the bias could very well be due to TOPEX.

6. Conclusions

The USC/JPL GAIM model uses a first-principles physics model of the ionosphere and a band-limited Kalman filter to assimilate multiple types of ionospheric measurements. Although one could use 2-D and 3-D tomographic inversion techniques to "image" the ionosphere, formal data assimilation techniques are required to fully exploit the information in a physics-based model. Comparison of GAIM analysis and GIM shows very close agreement, which is a minimum requirement for a successful assimilation model given that the two approaches rely on the same data source (ground TEC). However, comparison of GAIM and GIM to TOPEX VTEC indicate that GAIM is superior to GIM in that it is able to capture sharp changes in VTEC much better than GIM, the latter being limited by the thin shell model and the lack of physics. Future GAIM work will include the validation of electron density and the simultaneous estimation of the ion densities and some of the ionospheric drivers by means of an extended Kalman filter.

Acknowledgements

We thank Graham Bailey and Arthur Richmond for their assistance in the effort of building GAIM. This work is supported by the Department of Defense through a Multiple-disciplinary University Research Initiative. The research conducted at the Jet Propulsion Laboratory is under a contract with NASA.

References

Bailey, G. J., R. Sellek, and Y. Rippeth, A modeling study of the equatorial topside ionosphere, *Ann. Geophy.,* 11, 263-272, 1993.

Bierman, G. J., *Factorization Methods for Discrete Sequential Estimation*, Academic Press, 1977.

Codrescu, M. V., K. L. Beierle, T. J. Fuller-Rowell, S. E. Palo, X. L. Zhang, More total electron content climatology from TOPEX/Poseidon measurements, *Radio Sci.*, **36** (2): pp. 325-333, 2001

Dymond, K. F., S. A. Budzien, A. C. Nicholas, S. E. Thonnard, R. P. McCoy, R. J. Thomas, Electron densities determined by inversion of ultraviolet limb profiles, *J. Geophys. Rev*,, **106** (A12): 30315-30321, 2001.

Fejer, B. G., E. R. de Puala, S. A. Gonzales, and R. F. Woodman, Average vertical and zonal *F*-region plasma drifts over Jicamarca, *J. Geophy. Res.,* **96**, 13901-13906, 1991.

Ghil, M. and P. Malanotte-Rizzoli, Data assimilation in meteorology and oceanography, *Adv. Geophys.*, **33**, pp. 141-266, 1991.

Ghil, M., K. Ide, A. F. Bennett, P. Courtier, M. Kimoto, and N. Sato (Eds.),: *Data Assimilation in Meteorology and Oceanography: Theory and Practice*, Meteorological Society of Japan and Universal Academy Press, Tokyo, 496 pp., 1997.

Hajj G. A., and L. J. Romans, Ionospheric electron density profiles obtained with the global positioning system: results from the GPS/MET experiment. *Radio Sci.,* **33**, 1, pp. 175-190, 1998.

Hajj G. A., L. C. Lee, X. Pi, L. J. Romans, W. S. Schreiner, P. R. Straus, C. Wang, "COSMIC GPS ionospheric sensing and space weather", Terrestrial, atmospheric and oceanic sciences, **11**, 1, pp 235-272, Mar. 2000.

Ho, C. M., B. D. Wilson, A. J. Mannucci, U. J. Lindqwister, D. N. Yuan, A comparative study of ionospheric total electron content measurements using global ionospheric maps of GPS, TOPEX radar, and the Bent model, *Radio Sci.,* **32** (4): 1499-1512, 1997.

Kalman, R., A new approach to linear filtering and prediction problems, *Trans. ASME, Ser. D, J. Basic Eng.,* **82**, pp. 35-45, 1960.

Kalman, R. and R. Bucy, New results in linear filtering and prediction theory, *Trans. ASME, Ser. D, J. Basic Eng.,* **83**, pp. 95-108, 1961.

Hedin, A. E., Extention of the MSIS thermospheric model into the middle and lower atmosphere, *J. Geophys. Res.,* **96**, pp. 1159-1172, 1991.

Hedin, A. E., et al., Empirical wind model for the upper, middle and lower atmosphere, *J. Atmos. Terr. Phys*., **58**, pp. 1421-1447, 1996.

Heppner, J. R., and N. C. Maynard, Empirical high-latitude electric-field models, *J. Geophys. Res.,* **92**, pp. 4467-4489, 1987.

Leitinger R, Ladreiter HP, Kirchengast G, Ionosphere tomography with data from satellite reception of Global Navigation Satellite System signals and ground reception of Navy Navigation Satellite System signals, *Radio Science*, **32**, 4, pp. 1657-1669, 1997.

Lorenc, A., A global three-dimensional multivariate statistical interpolation scheme, *Mon. Weather Rev.,* **109**, pp. 701-721, 1981.

Mannucci A., B. Wilson, D. Yuan, C. Ho, U. Lindqwister, and T. Runge, A global mapping technique for GPS-derived ionospheric total electron content measurements, *Radio Science*, **33**, 3, 565-582, 1998.

Hernandez-Pajares, R. M., J. M. Juan, J. Sanz, M. Garcia-Fernandez, Performance of different TEC models to provide GPS ionospheric corrections, *J. Atmos. And Solar-Terr. Phys*. **64**, 18, pp. 2055-2062, 2002.

Pi, X., A. J. Mannucci, U. J. Lindqwister, and C. M. Ho, Monitoring of global ionospheric irregularities using the worldwide GPS network, *Geophy. Res. Lett.*, **24**, 18, pp. 2283-2286, 1997.

Pi, X., C. Wang, G. A. Hajj, I. G. Rosen, B. D. Wilson, and Graham Bailey, Estimation of ExB drift using a global assimilative ionospheric model: an observation system simulation experiment, *J. Geophys. Res.,* to appear.

Roger D., *Atmospheric Data Analysis*, Cambridge University Press, 1991.

Rosen, I. G., C. Wang, G. Hajj, X. Pi and B. Wilson, An adjoint method based approach to data assimilation for a distributed parameter model for the ionosphere, *Proceedings of the 40th IEEE Conference on Decision and Control,* Orlando, Florida, Dec. 4-7, 2001.

Schreiner, W. S., S. V. Sokolovskiy, C. Rocken, D. C. Hunt, Analysis and avalidation of GPS/MET radio occultation data, *Radio Sci.,* **34**, pp. 949-966, 1999.

Tobiska, W. K., Revised solar extreme ultraviolet flux model, *J. Atmos. Terr. Phys.,* **53**, 1005-1018, 1991.

Wang, C., G. Hajj, X. Pi, I. G. Rosen, and B. Wilson, A Review of the Development of a Global Assimilative Ionospheric Model, Proceedings of Ionosphere Effects Symposium, Alexandria, VA, May 2002.

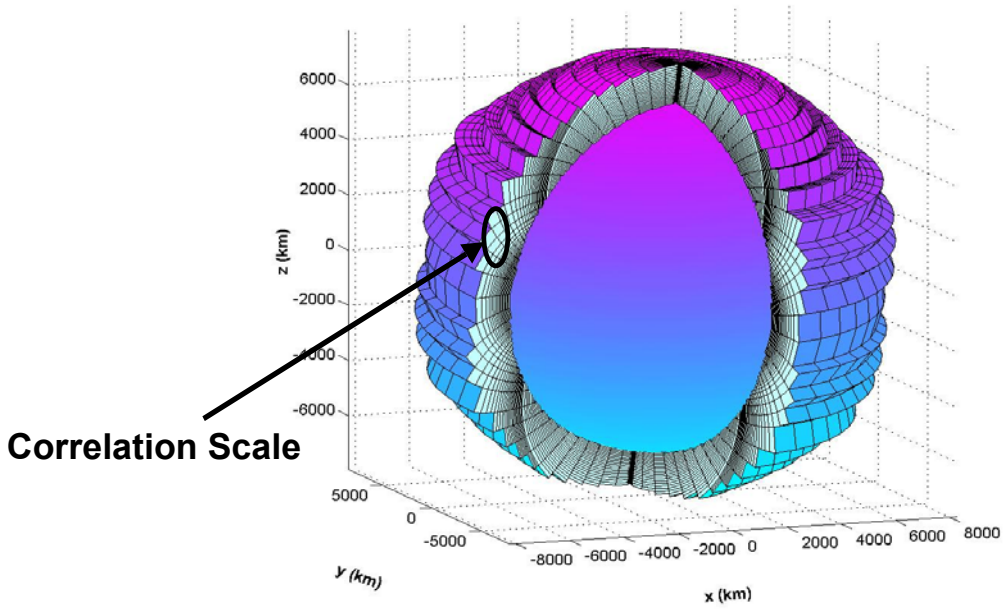


Figure 1. The grid used in modeling the ionosphere representing an Eulerian frame divided along constant geomagnetic field lines, constant geomagnetic potential lines, and constant geomagnetic longitudes. The ellipsoid represents the "correlation volume" used in setting the correlation between neighboring elements for the band-limited Kalman. An element centered at the ellipsoid will have zero covariance with elements outside the volume.

	1	2	3	4	5	6	7	8	9	10	11	12	...
1	1	0.5		0.5						0.5			
2	0.5	1	0.5		0.5						0.5		
3		0.5	1			0.5						0.5	
4	0.5			1	0.5		0.5						
5		0.5		0.5	1	0.5		0.5					
6			0.5		0.5	1			0.5				
7				0.5			1	0.5					
8					0.5		0.5	1	0.5				
9						0.5		0.5	1				
10	0.5									1	0.5		
11		0.5								0.5	1	0.5	
12			0.5								0.5	1	
:													

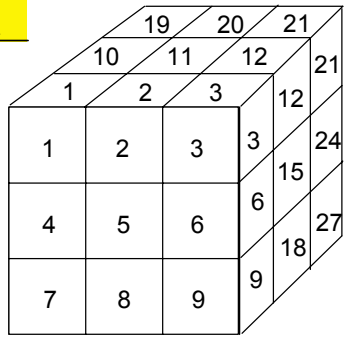


Figure 2. An example of a covariance matrix for a 3-D structure with non-zero correlation between immediate neighbors.

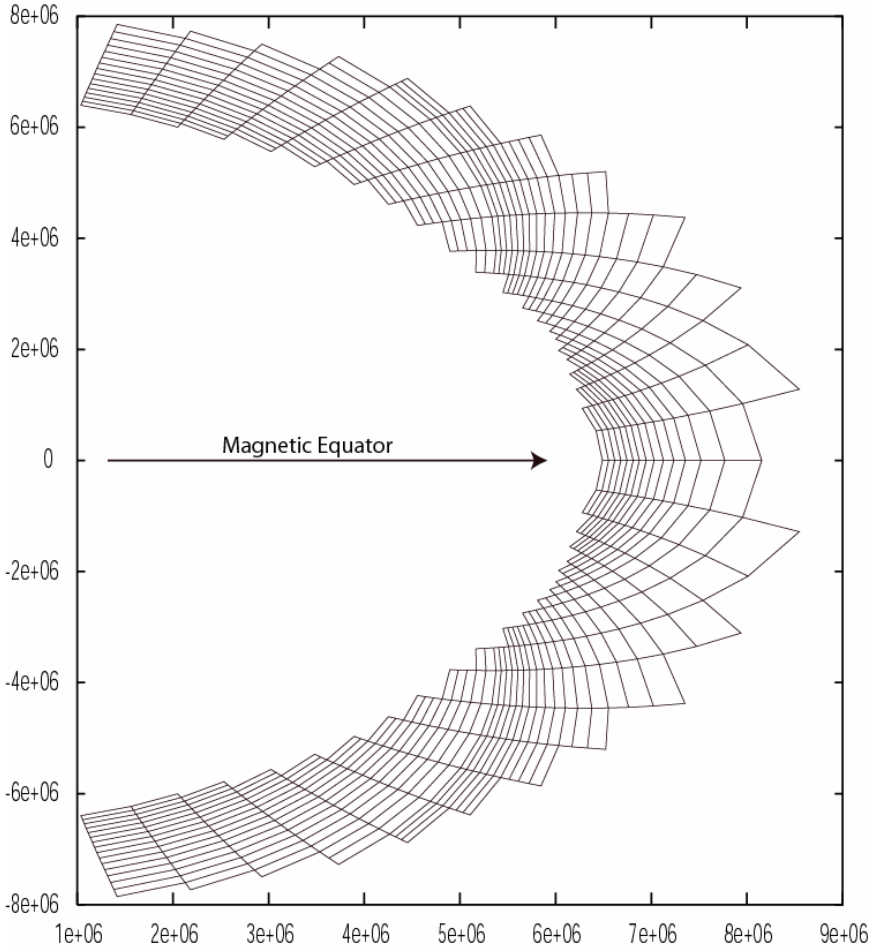


Figure 3: An example of a cross section at one magnetic longitude of the volume elements used in GAIM. These are defined by intesecting constant magnetic field lines, constant magnetic geopotential lines and constant magnetic longitudes. The vertical axis is alligned with the magnetic dipole used to model the Earth's magnetic field.

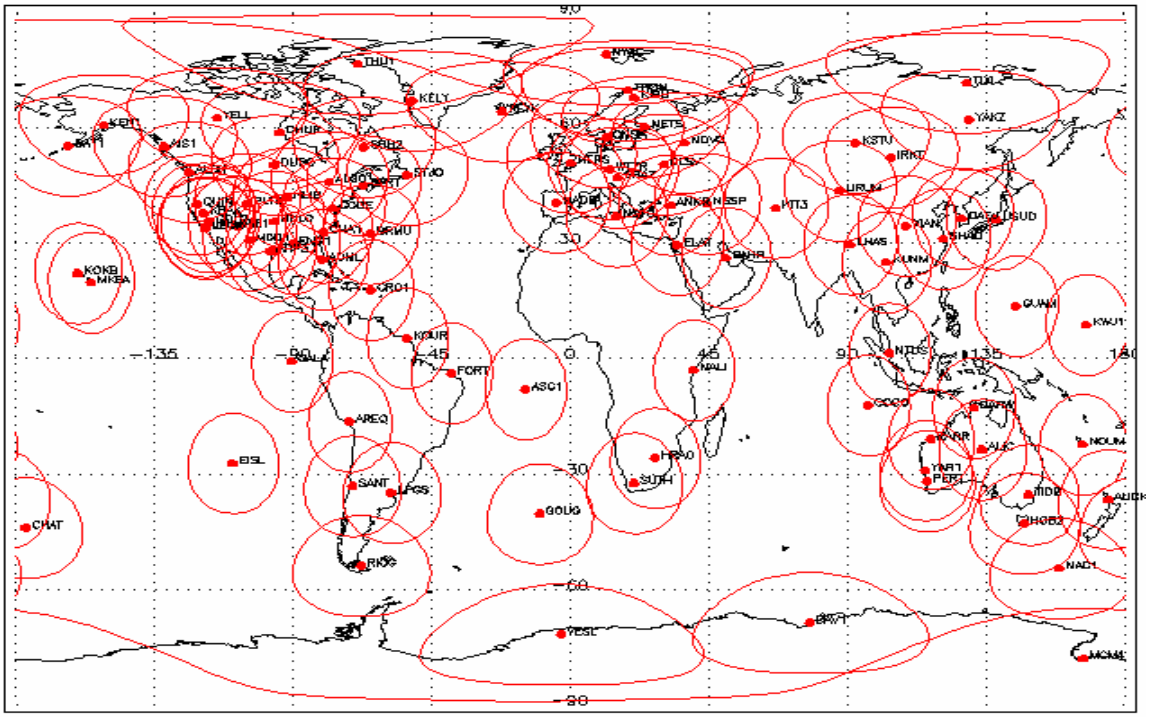
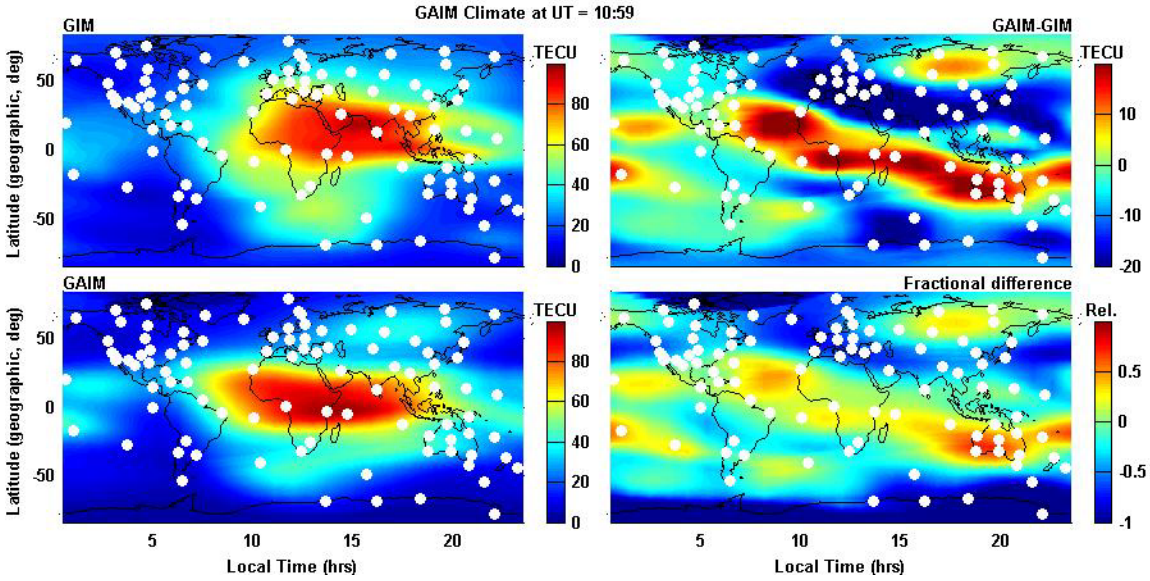
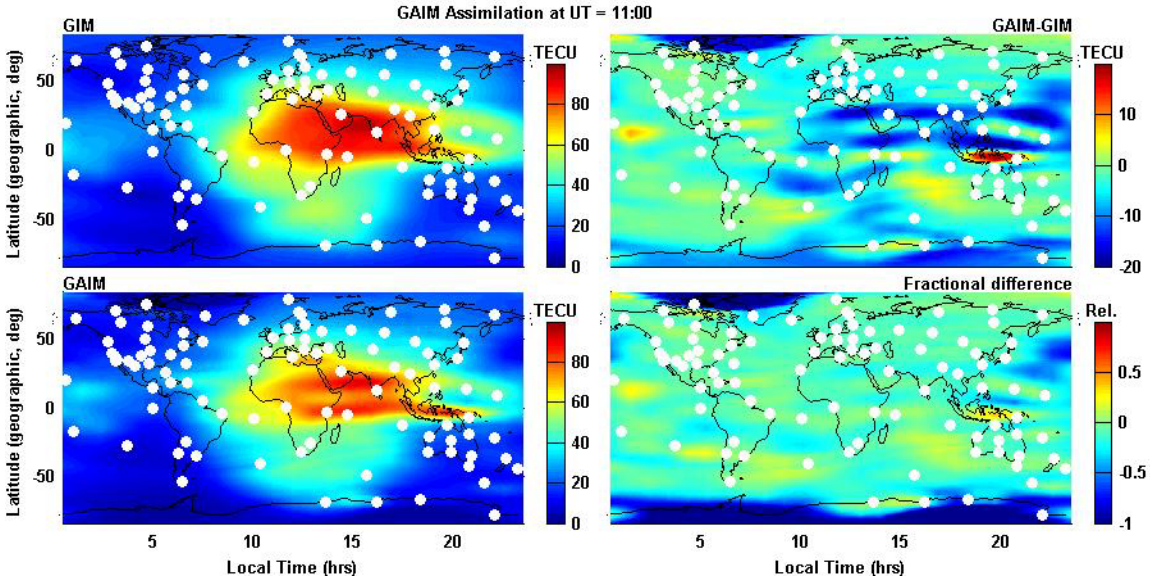


Figure 4: Coverage from the 98 ground GPS stations used in GAIM. The contour around each station corresponds to the visible region of the ionosphere at 450 km height for a 10 degree elevation mask.



(a)



(b)

Figure 5: (a) Comparison of GIM and GAIM climatology at about 1100 UT on May 23, 2002. Top-left: Global Ionospheric Map (GIM) of vertical TEC. Bottom-left: global vertical TEC obtained from vertically integrating GAIM climatology runs. Top-right: the difference between GAIM climatology and GIM. Bottom-right: fractional difference of GAIM climatology and GIM (defined as 2[GAIM-GIM]/[GAIM+GIM]). (b) Same as (a) but showing the GAIM analysis instead of the GAIM climatology. The dots indicate the GPS ground receivers. The GAIM analysis shows the equatorial anomaly more distinctly than GIM.

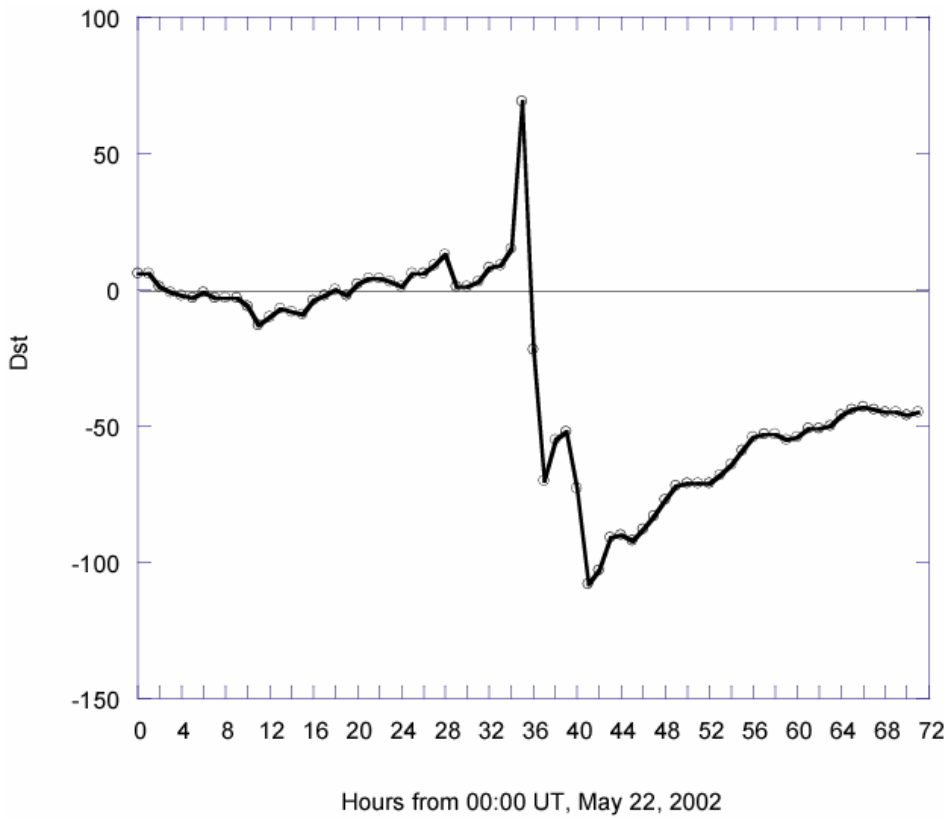


Figure 6: Dst values for the period May 22-24, 2002 showing the main phase of a storm between 13:00 and 17:00 UT on May 23rd.

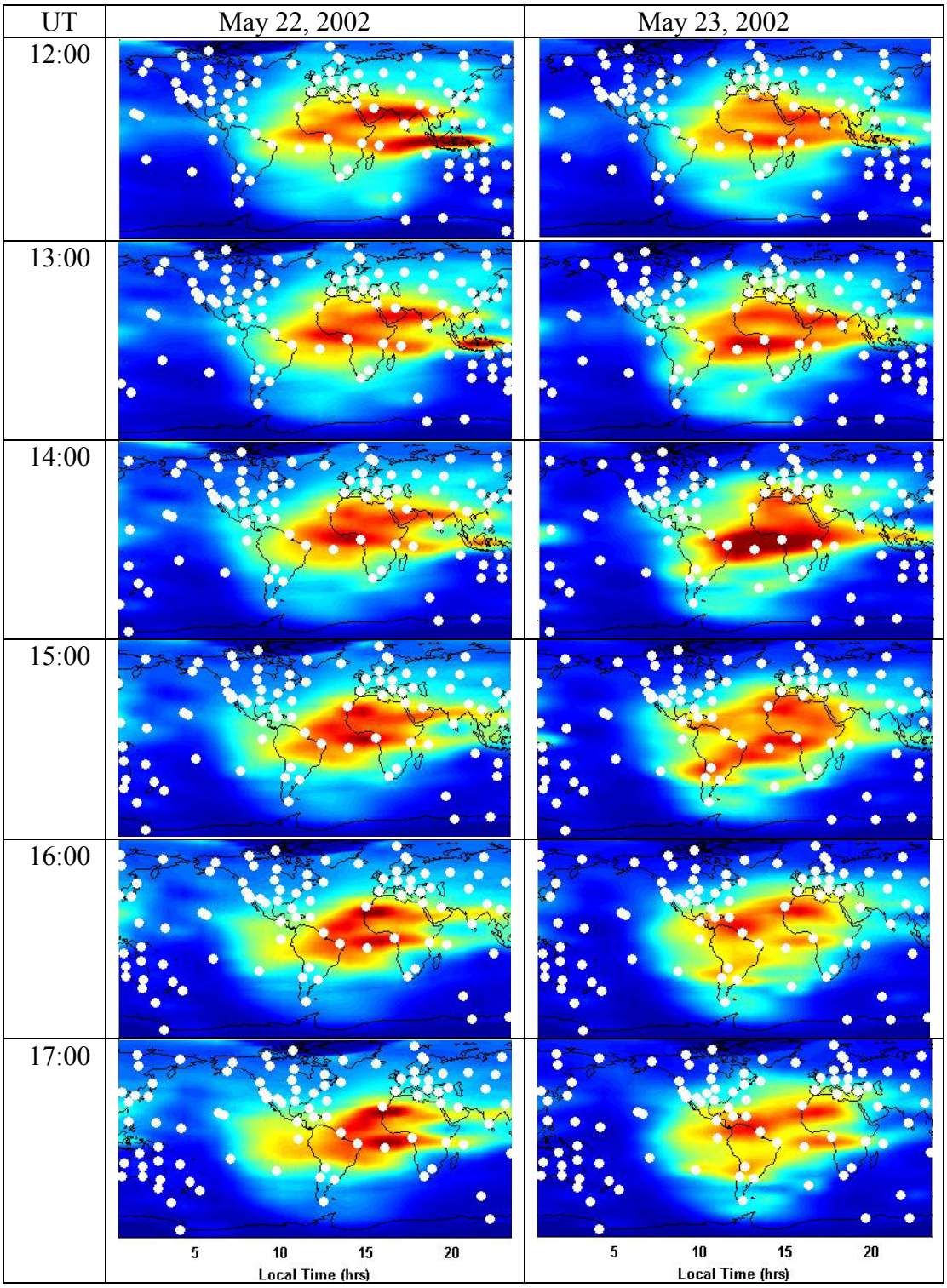


Figure 7: Hourly global maps of vertical TEC obtained by integrating hourly 3-D GAIM analyses for the days indicated on the first row and the UT indicated on the first column. A disturbed day (23rd) is shown next to a quiet day (22nd) for comparison.

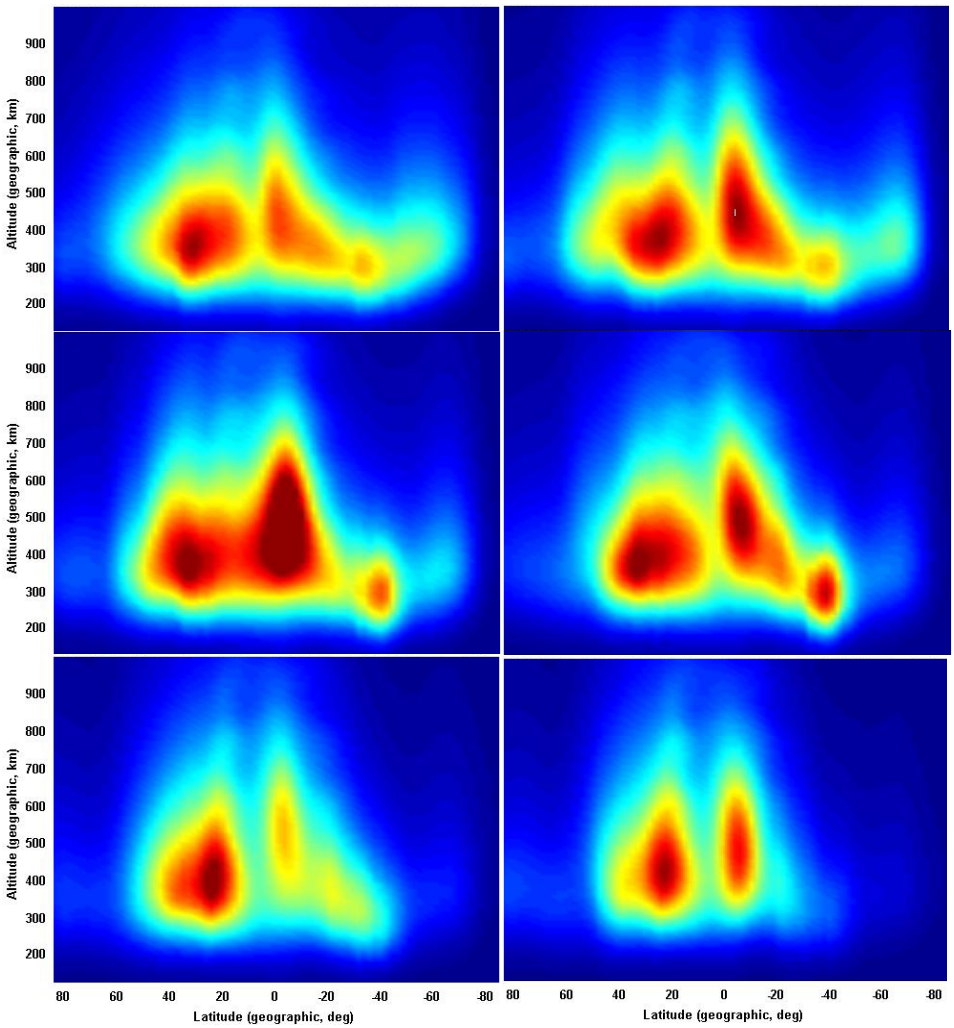


Figure 7: Hourly electron density snapshots at 7.5 geographic deg. longitudinal planes obtained from the same hourly GAIM analyses shown in Figure 6 for May 23, 2002. From left to right, top to bottom, the UT for each snap shot is 12:00, 13:00, …, 17:00, respectively. The local time is given by UT+30 minutes. The color scale ranges from 0 (blue) to 2×10^{12} e/m^3 (dark red) with some regions indicating over saturation.

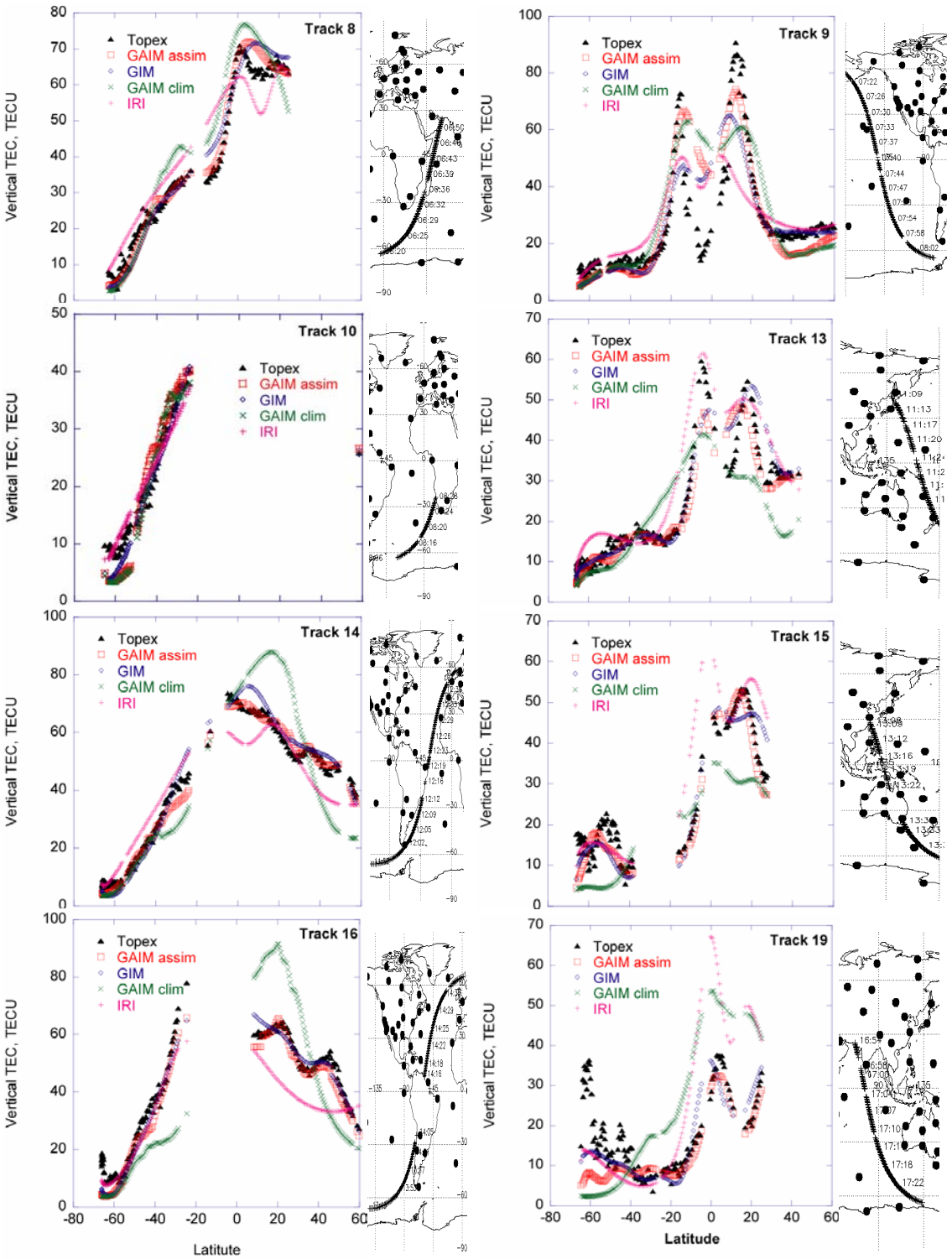


Figure 8: Comparison of vertical TEC below the TOPEX track for different tracks on May 23, 2002. To the right of each figure is the TOPEX ground marked by UT and neighboring ground GPS receivers used in the assimilation. The left panels correspond to ascending TOPEX tracks with an ascending node at ~10 AM local time. The right panels correspond to descending tracks at ~10 PM local time.

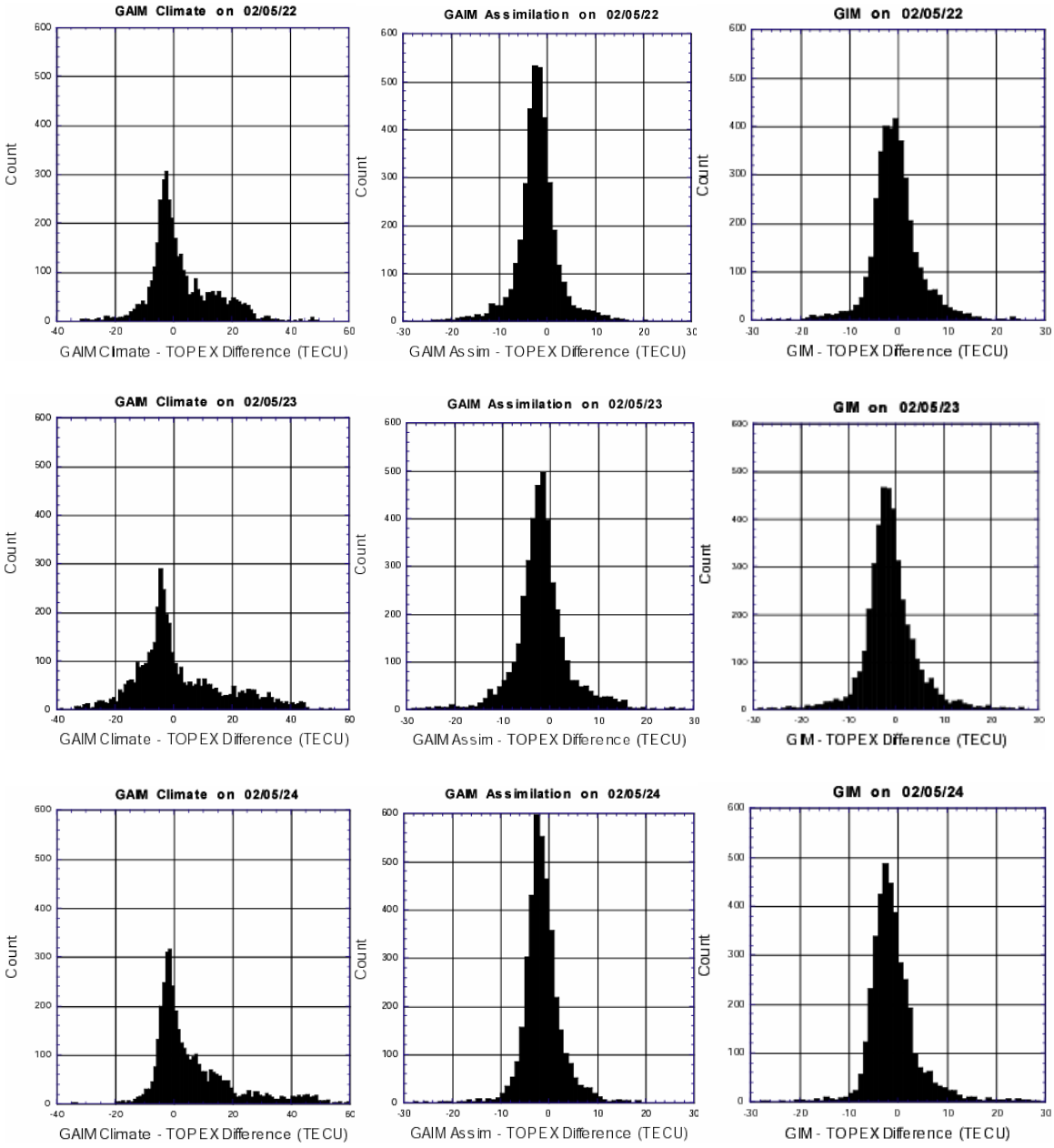


Figure 9: Histograms of vertical TEC differences between GAIM climatology, GAIM assimilation (i.e., analysis), or GIM and those obtained from TOPEX for the three days May 22-24, 2002.

Table 1. Specifications of the grid and correlation used in the data assimilation run.

Modeled region longitude range	0-360 deg.
Modeled region latitude range	-85 to 85 N.
Modeled region altitude range	100-1500 km altitude
Latitude resolution	5 deg.
Longitude resolution	15 deg.
Altitude resolution	80 km
Total no. of volume elements (voxels)	13,107
Correlation length in latitude	5 deg.
Correlation length in longitude	15 deg.
Correlation length in height	80 km

Table 2. Specifications of the physics input for the three days on May, 2002

Date	F10.7	Ap
May 22, 2002	185.6	8
May 23, 2002	184.8	78
May 24, 2002	193.9	2

Table 3. 3-hourly ap index on May 23, 2002

Time, in UTC	ap	Kp
01:30	12	3-
04:30	12	3-
07:30	7	2
10:30	111	7-
13:30	179	8-
16:30	236	8+
19:30	48	5
22:30	18	3+

Table 4: Statistics on vertical TEC differences between GAIM climatology (GAIM/C), GAIM analysis (GAIM/A), GIM, or IRI and those of TOPEX for the three days of May 22-24, 2002. The statistics include the total number of points for each day (N), the average, the standard deviation, the minimum and maximum differences and the RMS.

Day	Quantity	N	Ave	SD	Min	Max	RMS
22	GAIM/C –TOPEX	3763	2.3	10.8	-32	49.2	11
22	GAIM/A – TOPEX	3763	-2.0	4.9	-35	41.7	5.3
22	GIM –TOPEX	3763	-0.6	5.2	-30.7	28	5.2
22	IRI – TOPEX	3763	0.8	8.3	-38.2	32.5	8.3
23	GAIM/C –TOPEX	4102	1.6	15.3	-45.5	55.8	15.4
23	GAIM/A – TOPEX	4102	-1.7	6.1	-30.4	35.1	6.3
23	GIM –TOPEX	4102	-1.2	5.7	-35	29.3	5.8
23	IRI – TOPEX	4102	2.2	11.1	-48.2	56.4	11.3
24	GAIM/C –TOPEX	3901	6.4	14.2	-35	63.9	15.6
24	GAIM/A – TOPEX	3901	-1.3	4.4	-28.6	36	4.6
24	GIM –TOPEX	3901	-1.0	5.9	-36.7	36	6.0
24	IRI – TOPEX	3901	3.3	8.8	-39.5	38.1	9.4